SHREC'08 Entry: Visual Based 3D CAD Retrieval using Fourier Mellin Transform

Xiaolan Li[1,2]* Afzal Godil[1] Asim Wagan[1]

[1]National Institute of Standard and Technology, USA

[2]Zhejiang Gongshang University, China

ABSTRACT

Fourier Mellin Transform (FMT) has been used effectively in previous work for 2D image analysis, reconstruction and retrieval. In this paper, we perform 3D shape retrieval based on FMT on the Purdue Shape Benchmark. The whole procedure includes three steps: 1) generate silhouettes along the six principle directions for each 3D model; 2) compute a collection of FMT coefficients for all the silhouettes, which are translation, scale, and rotation invariant; and 3) compute a match measure between the query coefficients collection and those in the 3D shape repositories. The main contribution of this paper is the novel approach to extract the 3D signatures by Fourier Mellin Transform. Our experimental results validate the effectiveness of our approach.

KEYWORDS: Fourier Mellin Transform, visual similarity-based retrieval.

INDEX TERMS: J.6.1 [Computer-aided Engineering]: Computer aided design— [I.5.4]: Pattern Recognition —Applications

1 INTRODUCTION

3D models are widely used in several fields, such as computer graphics, computer aided design, computer aided manufacturing, molecular biology, and culture heritage, etc. The developments in 3D scanners and computer aided modeling software make it possible to quickly construct complete 3D geometry models with relatively low cost, which has triggered the rapid enlargement of 3D shape repositories. This has given impetus to research in the area of 3D content based retrieval. Several methods have been proposed to solve the problem, among which, except the hybrid method, visual similarity-based methods [1] are superior to the others [5]. Based on the above conclusions, this paper proposes a novel visual similarity-based approach for 3D shape retrieval.

Most of the former visual similarity-based approaches use the projected silhouette images or outline images of the shape, and then extract their image metrics, which are robust against transformations such as translation, rotation, scaling, and image distortion. Few methods start from depth buffer images captured from different directions of the model [6], which preserve 3D depth properties and are supposed to be more discriminating than the methods based on silhouettes. In this paper, Fourier Mellin Transform is applied on silhouettes to generate signatures to represent 3D models, because of its good performance in 2D image reconstruction, and retrieval. This is fundamentally different from most existing approaches for 3D shape searching.

We start by capture images from six principal directions of the model. The principal directions are generated by the PCA transform, which normalizes the orientation of the model. Using Fourier Mellin Transform, similarity transform invariant 2D features are extracted from those images. The concatenation of these 2D features is used to represent the 3D model. After that, the constraint rotation comparison is done between query model and the other models in the database to get the best retrieval result.

The rest of this paper is organized as follows. Section 2 describes some of the related work to our proposed approach for 3D object retrieval. The procedures of feature extraction and distance computation are described in Section 3 and 4 respectively. In Section 5, we provide the 3D model retrieval results on Purdue CAD Benchmark and some conclusions.

2 RELATED WORK

Designing discriminating 3D shape signatures is an active research area. Without being complete, we classify 3D shape searching techniques based on the shape representation into three categories: (a) visual similarity-based [1][3]; (b) geometric similarity-based [4]; (c) topologic similarity-based. Our approach is closely related to the visual similarity-based methods.

Orthogonal projected silhouette image is the most popular when researchers thinking about using the collection of 2D images to represent the 3D shape. Chen et al. [1] first captured 100 silhouettes with 10 different configurations of cameras mounted on 10 dodecahedrons. Then 35 Zernike moments coefficients and 10 Fourier coefficients calculated from one image are concatenated as one descriptor. After that, the similarities between objects are gotten with a special defined comparison. Vranic [6] only use 6 silhouettes, which is comparably less than that in [1], because of the pre-align procedure (PCA). Except the silhouette, grayscale depth image [6] is also a prevalent choice. Ohbuchi et al [3] extracted 42 generic Fourier descriptors for depth buffer images, which were captured from 42 different views. Again, after PCA align procedure, Vranic [6] only used 6 depth buffer images from principal directions to derive the shape descriptors by Fourier transform. The following course was the one to one comparison using the descriptors derived before.

From the above paragraphs, we come to the conclusion that the key technical point for a visual similarity-based approach performing effectively is to represent the images with similarity transform invariant descriptors, such as geometric moments, complex moments, Legendre moments, Zernike moments and Fourier descriptors etc. Fourier and Zernike are superior to the others according to the research of Zhang et al [7], which are a contour shape descriptor and a region shape descriptor respectively. Nevertheless, they are not completely satisfying the invariant requirement with respect to similarity transformations (i.e. rotation, translation and scale), and few of them can deal with grayscale images. Under the Fourier-Mellin transform framework, complete invariant descriptors can be derived [2], which is widely used in image reconstruction and image retrieval.

In this paper, we investigate a new method for 3D shape retrieval based on orthogonal projected silhouettes under the Fourier-Mellin transform framework.

* e-mail: lixlan@nist.gov

IEEE International Conference on Shape Modeling and Applications 2008
4 - 6 June, Stony Brook, New York, USA
978-1-4244-2261-6/08/$25.00 ©2008 IEEE

3 FEATURE EXTRACTION

Before computing the descriptors, the models need to be normalized with respect to orientation and translation [6].

Based on the modified PCA method, each 3D model is rendered using parallel projection with silhouettes from six principle directions. The procedure of Fourier-Mellin transform is described below.

1) A 2D FFT

$$F(u,v)=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f(x,y)e^{-j2\pi(ux+vy)}dxdy$$

is applied to the images captured before;

2) A log polar transform is performed to the images composed of the magnitudes of the Fourier coefficients;

3) Another 2D FFT is carried out on the log polar images to obtain the Fourier Mellin coefficients;

Because of the symmetry property of 2D FFT, we choose the magnitudes of the Fourier Mellin coefficients located in the first quadrant as the descriptors. Concatenating the descriptors acquired from six orientations, we obtain the complete representation for a 3D model as $FV= (X^+, X^-, Y^+, Y^-, Z^+, Z^-,)$.

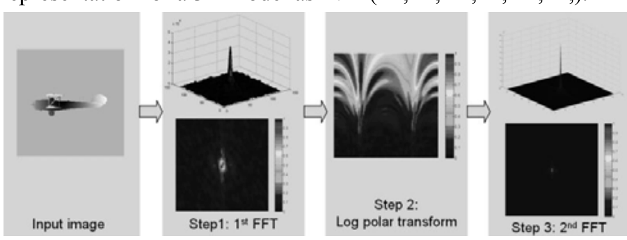

Figure 1. The FMT procedure.

4 DISTANCE COMPUTATION

Although modified PCA provides a canonical coordinate system, it can not define the direction of axis, which means the positive axis and the negative axis can be flipped randomly. Apart from that, PCA is prone to be influenced by the noises, which will cause the instability of the whole coordinate system. In this case, x axis may become y axis or z axis, y axis may become z axis, or vise versa. These can be rectified by a constraint rotational comparison.

Assuming there are two coordinate systems, we record the stable one as S, the other one which can be rotated as R. For the S, the possible configurations of R are listed in Table 1, with respect to the constraint rotation.

For two specific shapes A and B, the constraint rotational comparison procedure is described as following:

1) Compute the distance d_0 between the descriptor A and B;

2) Rotate the axis of shape A according to the configurations in Table 1, and compute the distance d_k (k=1,…,N-1) between rotated descriptor A and the fixed descriptor B, where N=16;

3) Choose the minimum d_k (k=0,…,N-1) as the distance to represent the similarity of A and B.

Table 1. Possible configurations of R w.r.t. constraint rotations.

S	X+	X-	Y+	Y-	Z+	Z-
	X+	X-	Y+	Y-	Z+	Z-
	X+	X-	Z+	Z-	Y-	Y+
	X+	X-	Y-	Y+	Z-	Z+
	X+	X-	Z-	Z+	Y+	Y-
	X-	X+	Y+	Y-	Z-	Z+
	X-	X+	Z-	Z+	Y+	Y-
	X-	X+	Y-	Y+	Z+	Z-
R	X-	X+	Z+	Z-	Y+	Y-
	Z+	Z-	Y+	Y-	X-	X+
	Z-	Z+	Y+	Y-	X+	X-
	Z-	Z+	Y-	Y+	X-	X+
	Z+	Z-	Y-	Y+	X+	X-
	Y-	Y+	X+	X-	Z+	Z-
	Y+	Y-	X-	X+	Z+	Z-
	Y+	Y-	X+	X-	Z-	Z+
	Y-	Y+	X-	X+	Z-	Z+

5 EXPERIMENT RESULTS

We illustrate 3D shape retrieval results using the FMT descriptors extracted from silhouettes. Manhattan (L1) distance is chosen as the similarity measure. The size of the final descriptor is 64*64*6. 45 CAD models are used as query models, some of which belong to the 866 models in the Purdue CAD Benchmark, while some of them are not. Several retrieval examples are shown in Figure 2.

In this paper, we have used the Fourier Mellin Transform into 3D shape searching field. In order to alleviate the errors caused by PCA alignment, a constraint rotational distance computation are proposed with respect to the constraint of orthogonal coordinate system.

REFERENCES

[1] D.Y. Chen, M. Ouyoung, X.P. Tian, Y.T. Shen. On visual similarity based 3D model retrieval. *Computer Graphics Forum (Eurographics '03)*, 03: 223-232, 2003.

[2] S. Derrode, F. Ghorbel. Robust and efficient Fourier-Mellin transform approximations for gray-level image reconstruction and complete invariant description. *Computer Vision and Image Understanding*, 83(1): 57-78, July 2001.

[3] R. Ohbuchi, M.Nakazawa, T. Takei. Retrieving 3D shapes based on their appearance. *Proc. of the 5th ACM SIGMM international workshop on Multimedia information retrieval*, 39-45, 2003.

[4] R. Osada, T. Funkhouser, B. Chazelle, D. Dobkin. Shape distributions. *ACM Transaction on Graphics*, 21(4):807-832, 2002.

[5] P. Shilane, P. Min, M.Kazhdan, T. Funkhouser. The Princeton shape benchmark. *Proc. of the Shape Modeling International 2004 (SMI'04)*, 04(00): 167-178, 2004.

[6] D.V. Vranic. 3D model retrieval. *Ph. D. Disserta- tion*, 2003.

[7] D. S. Zhang, G. Lu. An integrated approach to shape based image retrieval. *Proc. of 5th Asian Conf. on Computer Vision (ACCV)*, 02: 652-657, January 2002.

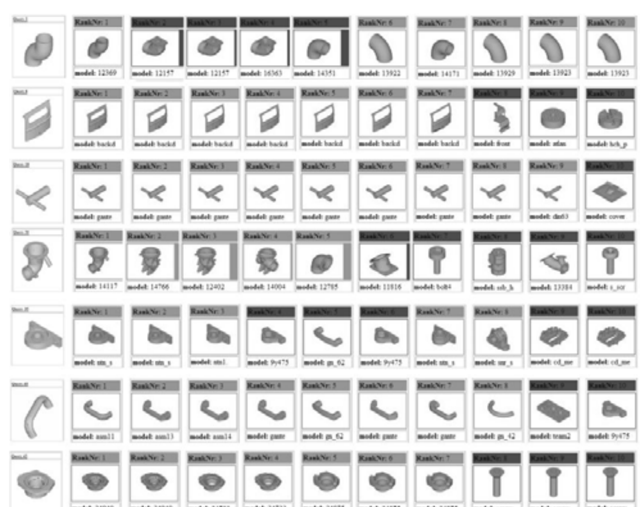

Figure 2. The retrieval results. The left one in each row is the query model. The following models in each row listed from left to right display the retrieval models rank 1 to rank 10, in which the irrelevant is highlighted with dark gray tag.